

19 October 1978

Nature Vol. 275 19 October 1978

617

NA strand of IS2
aligned to proceed
at position 33/52
cause of the high
night occasionally

occur (Fig. 7b). Subsequently inverted repeat pair 5 of the newly synthesised DNA strand could fold back on itself forming a new template for further DNA synthesis (Fig. 7c). The polymerase could then proceed in the 3' direction along the newly synthesised strand until it passes inverted repeat pair 3 and reaches position 33 (Fig. 7d). At this point the same slippage process is repeated: melting out, folding back of the inverted repeat pair 3 and DNA synthesis along the newly made DNA strand leading to the structure shown in Fig. 7e. This slippage tends to form long inverted repeats. Occasionally these may resist partial melting out so that DNA synthesis can then proceed along the old parental DNA strand, provided that inverted repeat pair 3 (33-52) in the parental DNA strand has folded up as indicated in Fig. 7e to yield a stable structure.

On another round of replication the DNA heteroduplex molecule forms two segregants, the old parental IS2 sequence and the newly generated IS2-6 allele.

The slippage of the DNA template during replication of IS2 as outlined in Fig. 7 generates symmetrical DNA additions (see Fig. 5). The process requires, first, a DNA region rich in A-T, and second, pairs of inverted repeat sequences at both ends of the A-T rich region. The newly formed sequences generated by this process may contain new signals for the turn-on of gene expression, as shown by IS2-6. Which parts of the IS2-6 sequences constitute the turn-on signal is not yet known.

The replication 'slippage' model described in Fig. 7 for IS2-6 allows us to predict the DNA sequence of the replication 'slippage' product of the upper DNA strand of IS2. Since preparation of this manuscript we have determined the sequence of the 54-base pair mini-insertion IS2-7. Its sequence is identical to the slippage product predicted for replication of the upper strand of IS2 and will be reported elsewhere¹³.

We thank Gudrun Hockema for technical assistance and Patricia Nevers for reading the manuscript. This work was supported by the Deutsche Forschungsgemeinschaft.

Received 8 May; accepted 4 August 1978.

1. Siedler, H. & Heiss, B. *Molec. gen. Genet.* 22, 267-277, (1973).
2. Starlinger, P. & Siedler, H. in *Curr. Topics Microbiol. Immun.* 15, 111 (1975).
3. Siedler, H., Reif, H. J., Hu, S. & Davidson, N. *Molec. gen. Genet.* 232, 263-289 (1974).
4. Pilchinsky, W. *et al. Gene* 2, 61-74 (1977).
5. Ghosal, D. & Siedler, H. *Molec. gen. Genet.* 158, 123-128 (1977).
6. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* 74, 560-564 (1977).
7. Musso, R. E. & Rosenberg, M. in *DNA Insertion Elements, Plasmids, and Episomes* (eds Bukhari, A. L., Shapiro, J. A. & Adhya, S. L.) 597-606 (Cold Spring Harbor, 1977).
8. Strewe, R. E. & Hobson, G. *Eur. J. Biochem.* 57, 595-606 (1975).
9. Thuring, R. W. L., Sanders, I. P. M. & Bort, P. *Analys. Biochem.* 66, 213-221 (1975).
10. Brack, Ch., Eberle, H., Bickle, T. A. & Yuan, R. *J. molec. Biol.* 10, 305-309 (1976).
11. Musso, R., Di Laura, R., Rosenberg, M. & de Crombrughe, B. *Proc. natn. Acad. Sci. U.S.A.* 74, 106-110 (1977).
12. Schlesinger, D. H., Niam, H. D. & Wilson, D. *Biochem. biophys. Res. Comm.* 61, 282-289 (1974).
13. Ghosal, D. & Siedler, H. *Cold Spring Harb. Symp. quant. Biol.* 43, (in the press).

A strand of IS2
ted repeat pairs 3
rows throughout.
ding to the repli-
a black line. For
both parts of an
hed lines indicate
page of the DNA
roduplex formed
32-6. See text for

Phenotypic expression in *E. coli* of a DNA sequence coding for mouse dihydrofolate reductase

Annie C. Y. Chang^{*†§}, Jack H. Nunberg[‡], Randal J. Kaufman^{||}, Henry A. Erlich[†], Robert T. Schimke[‡] & Stanley N. Cohen^{*†¶}

Departments of Genetics^{*}, Medicine[†], Biological Sciences[‡], Medical Microbiology[§], and Pharmacology^{||}, Stanford University, Stanford, California 94305

The construction and analysis of bacterial plasmids that contain and phenotypically express a mammalian genetic sequence are described. Such plasmids specify a protein that has enzymatic properties, immunological reactivity and molecular size characteristic of the mouse dihydrofolate reductase, and render host cells resistant to the anti-metabolic drug trimethoprim.

SINCE the initial propagation of eukaryotic DNA in bacteria¹, several systems have been used to study the expression in *Escherichia coli* of DNA derived from higher organisms. Biological activity of genes from the lower eukaryotes, *Saccharomyces cerevisiae*^{2,3} and *Neurospora crassa*⁴, has been demonstrated using phenotypic selection for functions that complement mutationally inactivated homologous bacterial genes. Immunological reactivity with antibody made against human somatostatin was shown for a peptide fragment cleaved *in vitro* from a hybrid protein encoded in part by bacterial DNA and in part by a chemically synthesised somatostatin DNA

sequence⁵. Very recently, a protein containing amino acids of rat proinsulin was shown to be made by bacteria that carry a double-stranded complementary DNA (cDNA) transcript of pre-proinsulin mRNA⁶; in that instance, antigenic determinants for both insulin and the bacterial enzyme β -lactamase were detected on a fused peptide transported outside the cell. It is not known, however, whether the mammalian peptide components of such immunologically reactive hybrid proteins have functional biological activity.

Our approach to the study of mammalian gene expression in bacteria has been to generate a heterogeneous population of clones carrying a DNA sequence that codes for a selectable mammalian gene product, and then to select directly those bacteria in the population that phenotypically express the genetic sequence. The mammalian enzyme dihydrofolate reductase (DHFR), which catalyses the conversion of dihydrofolic acid to tetrahydrofolic acid, is especially suitable for this purpose. The mammalian DHFR has a much lower affinity for the antimetabolic drug, trimethoprim (Tp), than does the corresponding bacterial enzyme⁷. Thus, bacteria which biologically express mammalian DHFR activity are resistant to levels of trimethoprim that ordinarily inhibit growth.

When these studies were initiated, the only bacterial host approved for EK2 recombinant DNA experiments⁸ was *E. coli* K12 strain χ 1776 (ref. 9). As this strain is already resistant to

[†]To whom correspondence should be addressed at Room S337, Stanford University Medical Center, Stanford, California 94305.

high concentrations of trimethoprim because of its *thy*⁻ mutation¹⁰, direct selection of bacteria that synthesise the mammalian DHFR could not be carried out. Therefore, our initial studies used an *in situ* hybridisation procedure¹¹ to identify χ 1776 clones which carried a mouse DHFR cDNA sequence, with the expectation that a highly sensitive indirect radioimmunoassay¹² would be used to detect any clones that expressed antigenic determinants of the protein product encoded by the gene.

Construction and cloning of chimaeric plasmids containing a DHFR ds cDNA

Figure 1 summarises the experimental scheme used in this study. Partially purified mRNA containing DHFR sequences from methotrexate-resistant AT-3000 mouse cells¹³ was used in the preparation of double-strand (ds) cDNA by RNA-dependent DNA polymerase (reverse transcriptase) and DNA polymerase I (ref. 14 and Fig. 2). Homopolymeric deoxy-C 'tails' were added to the unfractionated cDNA by terminal deoxynucleotidyl transferase (Fig. 2) and the section of the gel containing the predominant (1,500 base pairs) tailed cDNA was eluted. The material recovered from the gel was annealed with an equimolar concentration of pBR322 plasmid DNA that had been cleaved in the β -lactamase gene by the *Pst*I endonuclease¹⁴ and treated with terminal transferase to add homopolymeric dG tails at the cleavage sites. The extent of annealing of DHFR ds cDNA with the plasmid vector was monitored by electron microscopy¹⁶ using circle formation as an indicator. *Pst*I sites are regenerated at both ends of the insert as a result of such recircularisation^{17,18}.

Constructed chimaeric plasmids were introduced into MnCl₂-treated¹⁹ *E. coli* K12 strain χ 1776 and tetracycline (Tc)-resistant transformants (yield ~30 colonies per ng DNA) were selected and tested separately for the presence of a DNA species complementary to a highly purified DHFR cDNA probe. About 40% of the Tc-resistant clones gave a positive reaction by an *in situ* hybridisation test¹⁷ (Table 1 and Fig. 3). Plasmids from 14 of the reacting clones were examined by gel electrophoresis and all were found to contain a single DNA insert approximately 1,500 base pairs in length. As both ends of the *Pst*I-cleaved pBR322 DNA receive homopolymeric dG tails, only those molecules that have acquired a dE-tailed cDNA insert or had escaped cleavage and/or dG tailing would be expected to re-circularise and transform; thus, we presume that most of the nonreacting

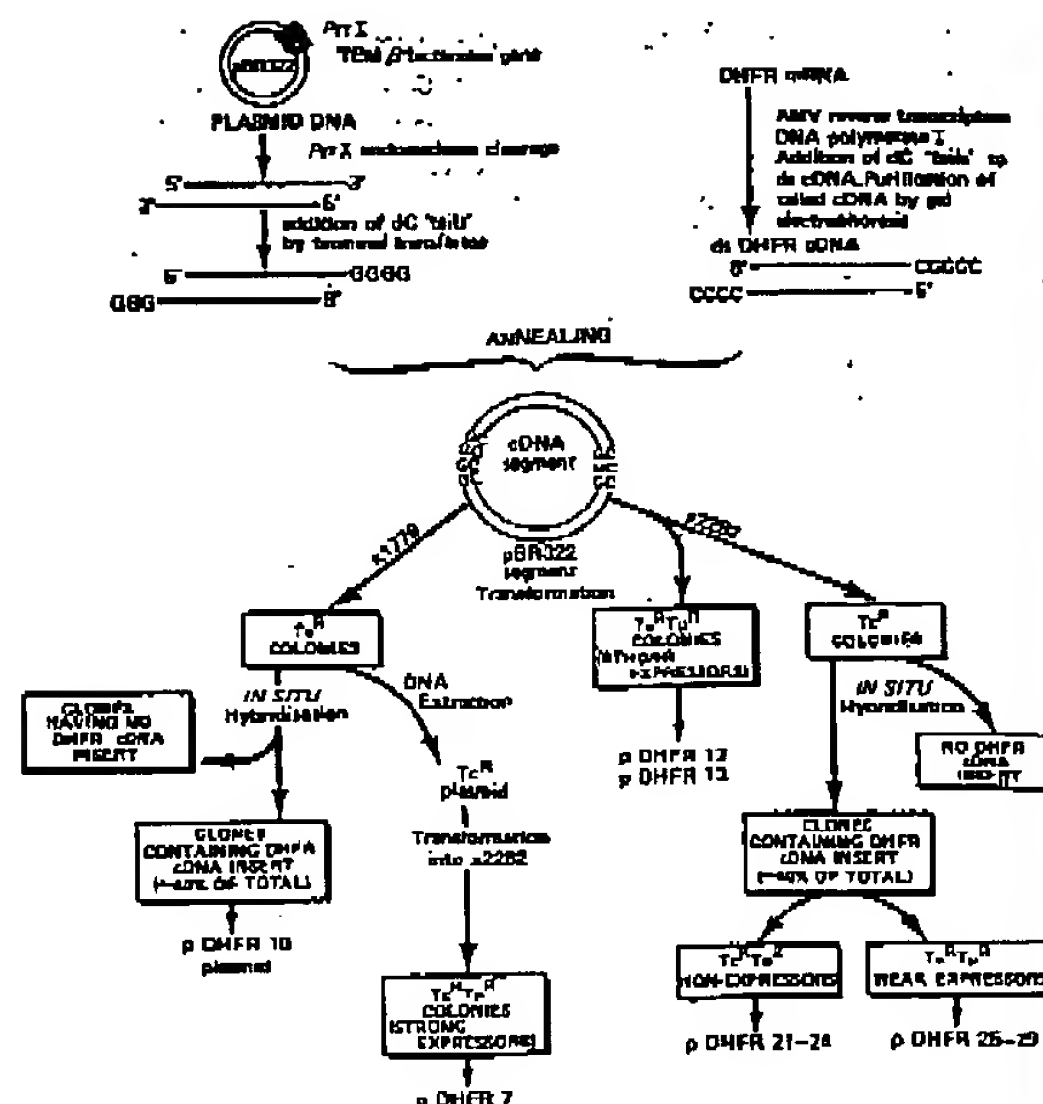


Fig. 1 Scheme used for cloning and expression of mouse DNA sequences that code for DHFR. Additional experimental details are given in the legend to Table 1 and in the text.

Tc-resistant colonies contain inserts of contaminating non-DHFR cDNA.

When χ 2282, a *thy*⁻ variant of χ 1776, was approved for EK2 use, direct selection of bacteria that phenotypically expressed the eukaryotic DNA sequence became feasible (see Fig. 1). Two populations (25 colonies each) of previously identified Tc-resistant colonies of χ 1776 were pooled, and plasmid DNA extracted from these populations was introduced by transformation into χ 2282. In other experiments, we transformed χ 2282 directly with annealed pBR322-DHFR cDNA (Table 1). Colonies that expressed resistance to both Tc and Tp were obtained in both types of experiments; three independent

Table 1 Transformation experiments using χ 1776 and χ 2282

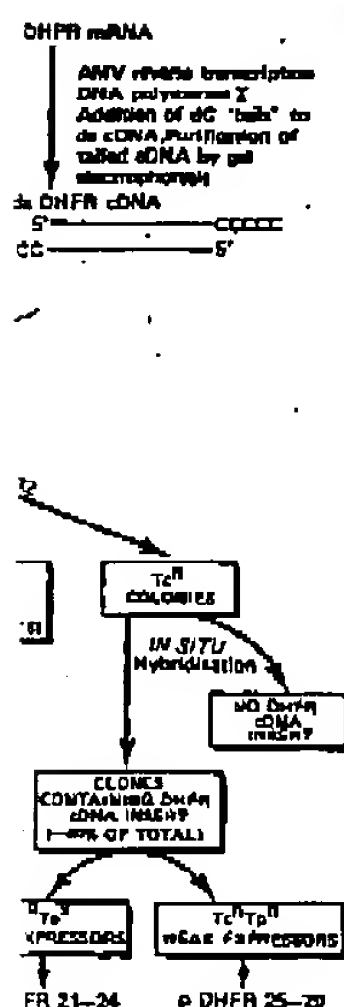
Host	Transforming DNA		Transformants per ng DNA			<i>In situ</i> hybridisation with DHFR cDNA probe (% positive)
	Plasmid vector	Insert	Tc (5 μ g ml ⁻¹)	Tc (10 μ g ml ⁻¹)	Tc/Tp (5 μ g ml ⁻¹ of each)	
χ 1776	pBR322	None	4 \times 10 ²	2 \times 10 ²	< 2 \times 10 ⁻³	—
χ 1776	pBR322	cDNA(1 st)	60	32	—	40
χ 2282	pBR322	cDNA(1 st)	70	—	2	44
χ 2282	pBR322	cDNA(2 nd)	60	—	1.3 \times 10 ⁻¹	—
χ 2282	pDHFR7	—	75	—	25	—

pBR322 plasmid DNA that had been annealed *in vitro* with dC-tailed DHFR cDNA (designated 1st) was introduced into χ 1776 or χ 2282, using a modification of a previously described transfection procedure¹⁹. 1 ml of an overnight bacterial culture was inoculated into 100 ml of L broth supplemented with diaminopimelic acid (DAP, 50 μ g ml⁻¹) and (for χ 1776 only) thymidine (4 μ g ml⁻¹). Bacterial cultures were grown until exponential phase at 35°C and then collected by centrifugation at 4°C. Cells were washed in 0.3 volume 10 mM NaCl, resuspended in 30 ml freshly prepared MCN buffer (70 mM MgCl₂; 40 mM sodium acetate, pH 5.6, and 30 mM calcium chloride) and chilled on ice for 20 min. Cells were collected, resuspended in 1 ml MCN and added in 200- μ l aliquots to 50 μ l DNA in TEN (10 mM Tris-HCl, pH 7.5, 0.1 mM EDTA, 50 mM NaCl) or MCN buffer. After chilling at 0°C for 30 min, reactions were incubated at 27°C for 5 min, chilled again for 30 min, and 50- μ l samples were plated on to Penassay broth agar supplemented with DAP, thymidine (for χ 1776), and antibiotics as indicated. When χ 2282 was used, the selective medium was M9 minimal agar supplemented with 0.5% casamino acids, biotin (2 μ g ml⁻¹), DAP (50 μ g ml⁻¹) and Tp (2.5–10 μ g ml⁻¹) plus tetracycline (Yc) or kanamycin (Km) as indicated. Plates containing transformants were incubated at 32°C and colonies were scored 2–3 d after plating. pBR322 plasmid DNA lacking the cDNA insert was used as a control. cDNA preparations labelled as 2nd consisted of plasmid DNA isolated from a nonfractionated population of clones that had previously been transformed with chimaeric molecules carrying a cDNA insert. In the experiment shown in the last line the transforming DNA was isolated from a clone (pDHFR7) that expresses resistance to Tp as well as Tc.

19 October 1978

Nature Vol. 275 19 October 1978

619



on of mouse DNA
xperimental details
in the text.

contaminating non-

s approved for EK2
typically expressed
ole (see Fig. 1). Two
usly identified Tc-
and plasmid DNA
roduced by trans-
ts, we transformed
R cDNA (Table 1).
Tc and Tp were
three independent

In situ
hybridisation with
DHFR cDNA probe
(% positive)

—
40
44
—

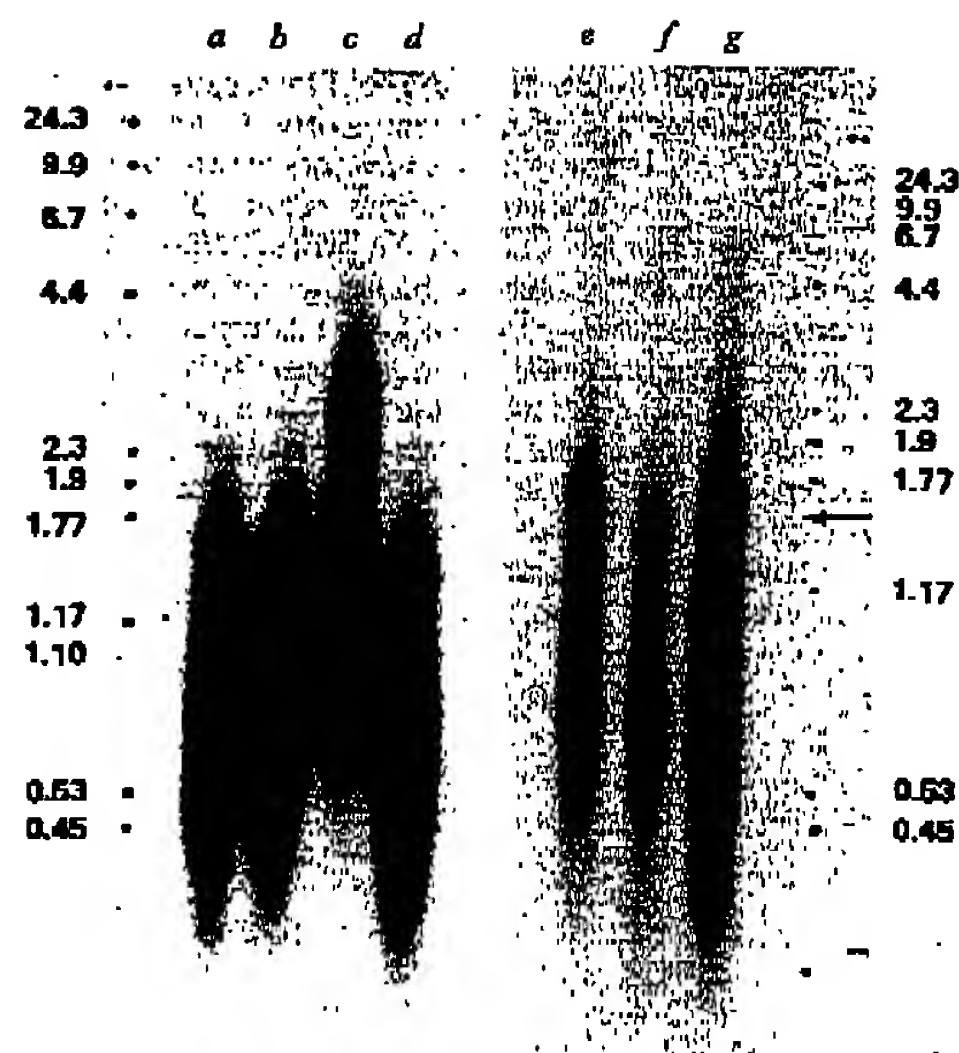
776 or χ 2282, using a
o 100 ml of L broth
es were grown until
nded in 30 ml freshly
Cells were collected,
mM NaCl) or MCN
les were plated on to
elective medium was
s tetracycline (Tc) or
ing. pBR322 plasmid
om a nonfractionated
shown in the last line

Fig. 2 Preparation and characterisation of DHFR cDNA insert. DNA complementary to DHFR mRNA was synthesised essentially as described elsewhere¹⁴, using avian myeloblastosis virus (AMV) reverse transcriptase and polysomal RNA obtained by indirect immunoprecipitation of DHFR-synthesising polysomes from methotrexate-resistant AT-3000 S-180 mouse cells¹⁵. The RNA had been estimated to contain DHFR mRNA as 20% of its mRNA¹⁷. The reaction was carried out in 100 μ l using 340 μ g polysomal RNA (estimated to contain 5 μ g polyA-RNA), 45 units AMV reverse transcriptase and dCTP labelled to 4 Ci mmol⁻¹ with ³²P-dCTP (Amersham). Approximately 1.4 μ g cDNA was synthesised in 30 min. The reaction was stopped by the addition of EDTA (to 10 mM) and extracted with phenol, followed by ether, before being passed over a Sephadex G-50 fine column in 10 mM Tris, pH 7.4, 2 mM EDTA and 10 mM NaCl (TEN). The void volume was collected and precipitated with ethanol. After centrifugation, the RNA was hydrolysed with NaOH (ref. 31), neutralised and precipitated with ethanol. The cDNA was then used as template for the synthesis of the second strand by *E. coli* DNA polymerase I essentially as described³¹. The reaction took place at 42 °C for 10 min in 100 μ l using 1.1 μ g cDNA, 10 units DNA polymerase I, and 200 μ M of each deoxynucleotide triphosphate with dCTP adjusted to 30 Ci mmol⁻¹ as above. Approximately 0.85 μ g of the second strand was synthesised. The reaction was stopped and extracted as above before being passed over a Sephadex G-50 fine column in TEN containing only 0.1 mM EDTA. Column fractions containing the ds cDNA were then treated with *Aspergillus oryzae* S₁ nuclease as described elsewhere³¹. After extraction and precipitation with ethanol, approximately 1.0 μ g ds cDNA was obtained. Aliquots of *a*, first strand product; *b*, first strand product after base treatment; *c*, second strand product and *d*, second strand product after S₁ nuclease treatment were examined on a 1.5% agarose gel in alkaline conditions as described³². Terminal addition of dCTP to the ds cDNA by terminal deoxynucleotidyl transferase (TdT), prepared as described elsewhere³³ was carried out by a modification³⁴ of the Co²⁺ procedure³⁵. The reaction was carried out in 500 μ l containing 140 mM cacodylic acid, 30 mM Tris base, 110 mM KOH (final pH 7.6), 0.1 mM dithiothreitol, 150 μ M dCTP (adjusted to 8 Ci mmol⁻¹ with ³H-dCTP (Amersham)), 1 mM CoCl₂ (added to prewarmed reaction mix before enzyme addition), approximately 1.0 μ g ds cDNA (assuming an average MW giving approximately 600 base pairs, this provides 10 pM 3' termini per ml) and 0.5 μ l TdT (2.3 \times 10⁵ units ml⁻¹). The reaction was allowed to proceed at 37 °C for 10 min before being cooled and sampled to determine incorporation. Approximately 30 dC residues were added per 3' terminus. The reaction was stopped, extracted, desalted and precipitated with ethanol as above. Aliquots of *e*, second strand product; *f*, second strand product after S₁ nuclease treatment, and *g*, dC-tailed ds cDNA were analysed on a 1.7% agarose gel in Tris-acetate-NaCl (ref. 36). The dC-tailed ds cDNA was then preparatively electrophoresed on a similar gel and the '1,500-base pair' region (arrow) cut out of the gel and electrophoretically eluted into a dialysis bag as described elsewhere³⁷. The eluted material was extracted as above, concentrated by lyophilisation and precipitated with ethanol. After centrifugation, the 1,500-base pair dC-tailed ds cDNA (approximately 80 ng) was redissolved in 10 mM Tris HCl, pH 7.4, 0.25 mM EDTA and 100 mM NaCl (annealing buffer). pBR322 plasmid DNA, isolated as described elsewhere³⁸ was digested with a 1.5-fold excess of *Pst*I endonuclease in conditions suggested by the vendor (New England Biolabs) and the linear plasmid DNA was cut out and eluted from a 0.7% agarose gel in TBE³⁹ as described above. The plasmid DNA was 'tailed' with dG residues using procedures similar to those described above. Approximately 15–20 dG residues were added per 3' terminus. Following extraction, the dG-tailed vector was passed over a Sephadex G-50 fine column in annealing buffer and the void volume was collected. Equimolar amounts of dC-tailed ds cDNA and dG-tailed vector DNA were allowed to anneal essentially as described by W. Rowe and R. A. Fratell (personal communication) except that the vector concentration was kept at 75 ng ml⁻¹ in the annealing reaction. Circularisation was monitored by electron microscopy¹⁸ and was typically about 20–40%. This annealed DNA was used directly for transformation into χ 1776 or χ 2282.

transformants were found to be capable of immediate growth in media containing at least 1,000 μ g ml⁻¹ of Tp, and were termed 'strong expressors'. Plasmid DNA isolated from these colonies was shown by repeat transformation to encode both the Tc and Tp resistance phenotypes. However, Tp^RTc^R transformants occurred in different experiments at only 20–60% of the frequency observed when selection of transformants was carried out for Tc resistance alone. Transfer of Tc-selected colonies on to minimal medium agar plates containing Tp showed that all such clones express resistance to at least 1,000 μ g ml⁻¹. Together, these findings suggest that phenotypic expression of Tp resistance may be delayed in transformants until a sufficient quantity of plasmid-specified DHFR has accumulated. Analogous results have been obtained with other antimicrobial drug-resistance determinants encoded by plasmids introduced into *E. coli* by transformation²⁰.

Structure of the DHFR cDNA

Gel electrophoresis of endonuclease-cleaved plasmid DNA from three separately derived bacterial clones that expressed high levels of trimethoprim resistance showed similar overall patterns. Using such data, a cleavage map (Fig. 4) of the cDNA insert of one of these (pDHFR7) was constructed.



Examination of the amino acid sequence of the mouse DHFR enzyme²¹ allowed us to assign one of the *Hae*III cleavage sites to the *trp-pro* (TGG-CCX) present at amino acid positions 24 and 25, thus localising the DHFR structural sequence to a position near the 5' end of the mRNA template used in synthesising the cloned cDNA. Other cleavage sites within the DHFR structural sequence were consistent with positions predicted by computer analysis of the amino acid sequence. DNA sequence analysis of two separate regions of the cDNA provided direct verification that the nucleotide sequence of the insert corresponds to the amino acid sequence reported for the mouse DHFR enzyme, and also confirmed that the coding sequence for DHFR is located at the ds cDNA equivalent of the 5' end of the mRNA.

The nucleotide sequence at the pBR322-cDNA junction nearest the 5' end of the mRNA used as template for DHFR cDNA (that is, at the *Pst*I site) is of special interest. The complement of the 'sense' strand of β -lactamase gene of the vector (J. G. Sutcliffe, personal communication) is interrupted at the *Pst*I site by a series of 11 dG residues added by the terminal transferase, and these are followed immediately by (1) an ATG (AUG) protein start codon, and (2) the codon for the first amino acid of the mouse DHFR structural gene. The sequence that codes for the mouse DHFR is in the same orientation as at that encoding the β -lactamase on the vector plasmid; however, the

number of incremental G residues (that is, 10) at the vector-insert junction ensures that the DHFR cDNA sequence is not in the same translational reading frame as the β -lactamase gene.

If the phenotypic expression we have observed for the mouse DHFR sequences in bacteria is the result of translational readthrough from signals that initiate protein chains within the β -lactamase gene, then the host bacterial cell must be able to circumvent the observed frame shift by a 'slippage' mechanism of translation. Such an event might potentially be aided by the long run of dG residues introduced at the pBR322-cDNA junction and could account for our observation (Table 1) that plasmids containing a DHFR cDNA insert yield more than twice as many expressors as would be expected from considerations of reading frame and orientation. However, the high level of functional expression observed for both primary and secondarily transformed clones of pDHFR7 does not seem to be readily explained by slippage of tRNA molecules during translation.

Perhaps a more likely explanation is that the PstI-polyG-ATG sequence that has been constructed preceding the coding sequence for DHFR serves as a binding and protein initiation site for the bacterial ribosome. Recent studies^{23,24} have identified sequences on mRNA in the 5' direction from the initiator codon that are complementary to the CCUCC sequence at the 3' end of the 16S ribosomal RNA species, proposed by Shine and Dalgarno²⁴ to be involved in the binding of mRNA to ribosomes. It is tempting to speculate that the mRNA transcript from the sequence at the pBR322-cDNA

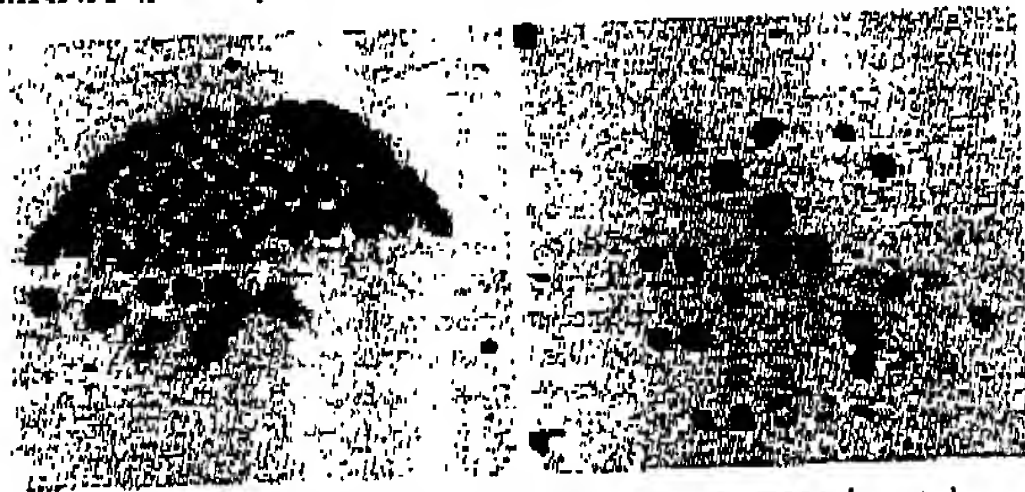


Fig. 3 Detection of colonies containing DHFR cDNA inserts by *in situ* hybridisation. Colonies were screened for DHFR sequences using a modification (G. N. Duell, unpublished) of an *in situ* hybridisation procedure¹¹. Tc-resistant colonies were transferred to nitrocellulose filters (Millipore, HAWG) that had been placed on Penassay broth agar plates containing Tc (10 μ g per ml). (Filters had been washed twice by boiling in H_2O and autoclaved before being placed on plates.) After 2-3 d of bacterial growth at 32°C, the filter was removed from the plate and placed on a Whatman no. 3 pad saturated with 0.5 M NaOH. After 7 min, the filter was sequentially transferred to a series of similar pads saturated with 1 M Tris, pH 7.4 (twice, 7 min each); 1.5 M NaCl, 0.5 M Tris, pH 7.5 (once, 7 min); and 0.30 M Na citrate (2 \times SSC) (once, 7 min). After the excess liquid had been removed by suction, the filter was placed on a pad containing 90% ethanol, dried by suction and baked *in vacuo* at 80°C for 2 h. Before hybridisation, filters were pretreated for 3-6 h at 65°C in hybridisation buffer that contained 5 \times SSC, pH 6.1 0.2% SDS, 0.02% Ficoll 400 (Pharmacia) and 8 μ g ml⁻¹ *E. coli* tRNA. Hybridisations were carried out with individual filters in 1.5 ml hybridisation buffer containing 2 $\times 10^4$ c.p.m. ³²P-labelled purified DHFR cDNA¹² in a sealed plastic bag at 65°C for 24 h. The filters were then washed in hybridisation buffer (once, 60 min at 65°C); in 5 \times SSC, pH 6.1 (three times, 60 min each at 65°C); and in 2 \times SSC, pH 7.4 (twice, 10 min at room temperature), air dried, and prepared for autoradiography. Left: top, a collection of χ 1776 colonies which contain a DHFR cDNA insert; middle, Tc-resistant χ 2282 colonies derived from transformation with annealed pBR322 ds cDNA—both reacting and non-reacting colonies are seen; bottom, colonies containing pBR322 and pACYC101 plasmids which show no visible hybridisation. Right: several positive colonies on a representative filter analysed in screening χ 1776 transformants. Negative colonies represent clones containing pBR322 or pACYC101 plasmids which show no visible hybridisation.

junction has sufficient complementarity to the CCUCC sequence to allow ribosomal binding when a translational start signal is located an appropriate distance away. In such an event, the ATG protein start signal that immediately precedes the coding sequence for the mouse DHFR might initiate a peptide chain having a size characteristic of the mammalian enzyme. Immunological analysis of extracts derived from pDHFR7 and other expressing clones has yielded results consistent with this interpretation (see below).

Analysis of enzyme activity encoded by pDHFR7 plasmid

Mammalian dihydrofolate reductases can be distinguished from their bacterial counterparts by the ability of mammalian enzymes to use folate as a substrate and by their differential sensitivity to competitive inhibitors^{25,26}. In initial experiments, the reduction of folate to tetrahydrofolate was measured using extracts from the pDHFR7 clone, from a Tp-sensitive clone containing a DHFR cDNA insert (the pDHFR10 plasmid), and from cells that contain only the pBR322 vector. Although all three clones are capable of synthesising a chromosomally produced bacterial enzyme, only the enzyme present in extracts from cells containing the pDHFR7 plasmid gave reduction of folate (4 \times background). Additional evidence that the reductase encoded by the pDHFR7 plasmid is of mammalian origin was obtained by inhibitor analysis (Fig. 5). The DHFR isolated directly from mouse cells and the activity encoded by the pDHFR7 plasmid showed identical sensitivities to methotrexate, trimethoprim and a triazine derivative (2,4,6-diamino-1-(4'-butylphenol)-6,6 dimethyl-1, 6-dihydro-1,3,5-triazine); both enzyme activities were 200 times more sensitive to the triazine than to trimethoprim. In contrast, bacterial dihydrofolate reductase is inhibited more effectively by trimethoprim ($K_i = 5 \times 10^{-9}$) than it is by triazine ($K_i = 6.5 \times 10^{-4}$)²⁵.

As methotrexate binds stoichiometrically to dihydrofolate reductase²⁷, we can estimate the number of molecules of enzyme in the pDHFR7 plasmid extracts from the methotrexate inhibition data of Fig. 5. We calculate from the specific activity of the extract (3 units per mg of soluble protein) and the specific activity and methotrexate binding parameters of the mouse enzyme²⁸ that 0.01% of the soluble bacterial protein is active mammalian DHFR.

Immunological characterisation of bacterial cell extracts containing mouse DHFR

Immunological evidence confirming the nature of the DHFR encoded by pDHFR7 and other plasmids that contain a mouse DHFR cDNA insert was obtained using a solid-phase sandwich radioimmunoassay¹². Tp^R clones of χ 2282 containing the independently derived plasmids pDHFR 7, 12 and 13 showed a strong reaction with rabbit antibody directed against mouse DHFR in an *in situ* immunoassay¹² (data not shown); protein that reacted with the antibody was also made by bacteria which showed low levels of phenotypic expression and by some clones that did not make a biologically functional DHFR (that is, were Tp sensitive). The nature of the antigen synthesised by Tp^S and Tp^R clones was examined more fully using a newly developed method (filter affinity transfer, or FAT procedure) for the *in situ* immunological characterisation of proteins in gels²⁹. This procedure depends on the covalent coupling of F(ab)₂ antibody fragments to a chemically derivatised and activated cellulose filter; antigen transferred on to the filter from an SDS-polyacrylamide gel is detected by subsequent incubations with antiserum and ¹²⁵I-labelled *Staphylococcus aureus* protein A (ref. 12).

Filter affinity transfer analysis of the pDHFR7 extract (Fig. 6, lanes b, c) shows the presence of protein that reacts immunologically with the antibody to mouse DHFR and further shows

621

622

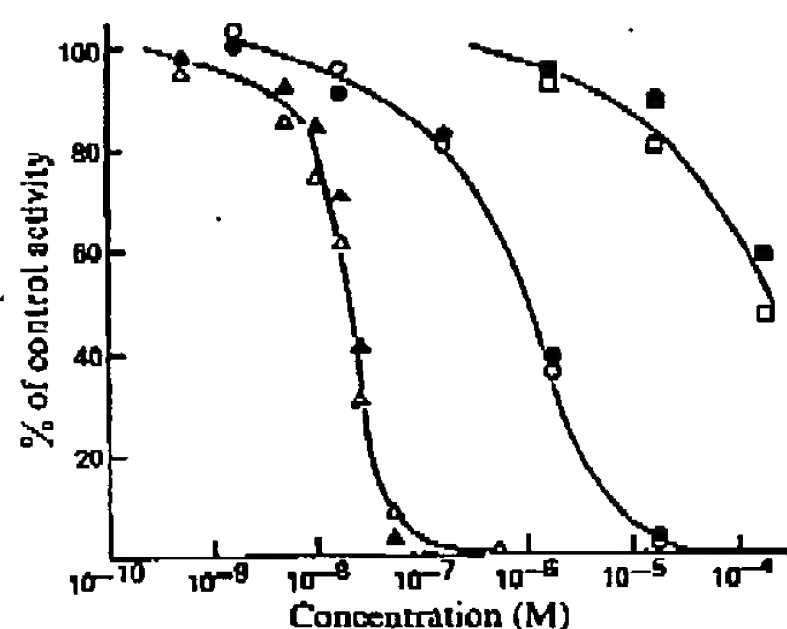


Fig. 5 Inhibitor analysis of DHFR from bacterial cells. Stationary phase cultures of $\chi 2282$ expressing trimethoprim resistance were grown in the presence of Tp ($1 \mu\text{g ml}^{-1}$) in minimal medium, washed with isotonic saline, and suspended in 50 mM potassium phosphate buffer, pH 7.0, containing 10 mM benzamidine and 10 mM phenyl methyl sulphonyl fluoride (3 volumes buffer to 1 volume cells). The suspension was sonicated and centrifuged at 10,000 r.p.m. for 15 min. The supernatant was centrifuged for 1 h at 100,000g before being studied. An R_2 methotrexate-resistant mouse cell extract was prepared as described elsewhere¹². Enzyme activity was measured by the radioactive folic acid assay previously described²⁸. Protein was determined by the method of Lowry⁴³. Approximately 3 units of activity from the $\chi 2282$ extract or 5 units from the methotrexate-resistant mouse cell extract were incubated with inhibitor for 10 min at 24°C before assaying for folate reductase activity; the concentrations shown represent the final concentration of inhibitor in the reaction mixture. Background values, determined by measuring enzyme activity in the presence of 10 mM methotrexate, have been subtracted from all points. The results presented are the average of duplicate samples which generally varied by less than 10% and are expressed as a percentage of the value obtained in the absence of inhibitor. One unit of activity is the amount of enzyme needed to reduce 1 nmol of folate in 15 min at 37°C. Δ and \blacktriangle indicate addition of methotrexate, \circ and \bullet indicate addition of the triazine derivative, and \square and \blacksquare indicate trimethoprim addition for $\chi 2282$ and the mouse cell extracts, respectively.

insert to the β -lactamase gene sequence of the vector, the length of each insert, and the minimal inhibitory concentration (MIC) determined for the clone are shown in Fig. 7. As can be seen, plasmids pDHFR7, 12, 13 and 26–29 all contain a complete DHFR structural sequence inserted in the same orientation (that is, orientation a) as the gene encoding the bacterial β -lactamase. The clone carrying each of these plasmids expresses Tp resistance, although the MIC varies from $150 \mu\text{g ml}^{-1}$ for pDHFR28 to $>1,000 \mu\text{g ml}^{-1}$ for pDHFR7, 12 and 13; the greatest reactivity with antibody to mouse DHFR occurs with pDHFR12 (unpublished data).

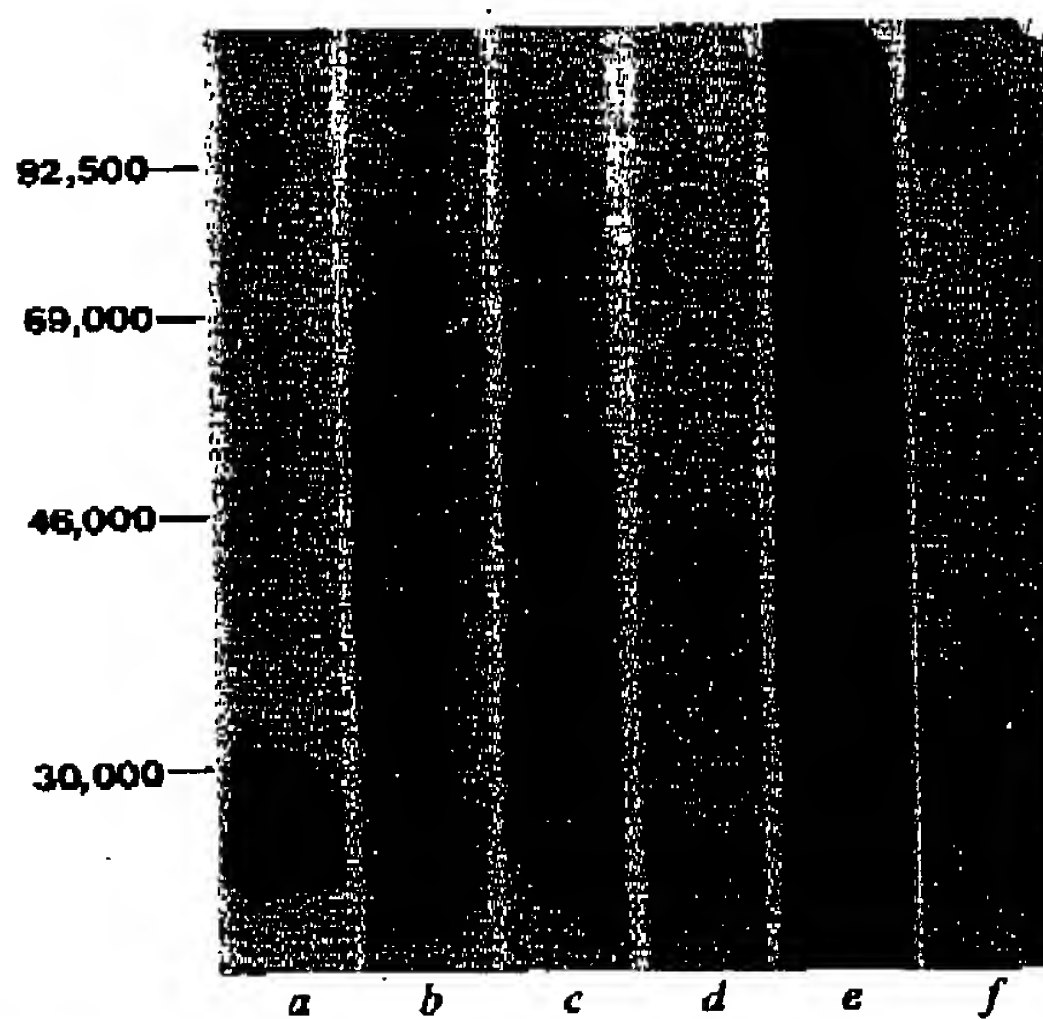
It is unlikely that translational reading frame is the determining factor in the different levels of expression observed in these clones, as our DNA sequence analysis indicates that a correct reading frame is not essential for efficient expression in pDHFR7, 12 or 13. However, the positioning of the putative ribosomal binding site in relation to the ATG start codon may potentially influence the strength of expression by affecting the formation of the translational initiation complex (compare with ref. 23) (Figs 4, 7).

The end of DHFR structural sequence that corresponds to the 5' end of the mRNA is not present in plasmids pDHFR 21 and 24, thus explaining the observed lack of functional expression of the cDNA in these clones. However, pDHFR21, which has lost less than 15% of the structural sequence, nevertheless encodes a (probably hybrid) peptide that contains antigenic sites which

react with antibody to mouse DHFR (Fig. 6 and unpublished data). It is particularly interesting that the clone carrying pDHFR25 expresses a low level of Tp resistance, although the coding sequence for the DHFR enzyme is inserted in an orientation opposite to that of the β -lactamase gene. This finding, and our detection in extracts of the pDHFR25 clone of protein that reacts immunologically with antibody to the mouse enzyme (unpublished data), suggest that readthrough transcription into the DHFR coding sequence from a promoter sequence located on or near the distal segment of the β -lactamase gene may occur. Consistent with this interpretation are preliminary data suggesting that DHFR antigenic sites are also synthesised by cells carrying pDHFR23, which is a non-expressor of Tp resistance and contains a cDNA insert in orientation b (Fig. 7). Further study is required to determine whether sequences in the distal segment of the β -lactamase gene are capable of serving as weak promoters for the initiation of mRNA chains that extend into the DHFR cDNA.

The findings reported here indicate that the bacterial clones we have constructed are synthesising and phenotypically expressing DHFR encoded by mouse cDNA sequences: (1) the cDNA insert cloned in bacteria has been shown by *in situ* hybridisation to be homologous with the mouse gene and by direct DNA sequence analysis to encode the amino acid sequence of mouse DHFR, (2) DHFR enzymatic activity and resistance to Tp are specified by nucleotide sequences present on chimaeric plasmids but not on the vector, (3) the DHFR

Fig. 6 Filter affinity transfer analysis²⁹ of bacterial cell extracts. 20 μl of extracts in SDS sample buffer were run at constant current for 3 h in an 11.25% SDS-polyacrylamide slab gel. The gel was incubated in PBS (50 mM phosphate buffer containing 0.15 M NaCl) for 30 min and placed on a blotter wet with PBS. Peptides were specifically transferred from the gel to strips of a dry cellulose filter that had been covalently coupled to anti-DHFR F(ab)₂ fragments¹². Filters were washed, incubated with antibody to DHFR, washed again and treated with ¹²⁵I-labelled protein A. After additional washing and drying steps, the filters were analysed by autoradiography. The eluate fraction (lane d) was obtained by passage of an extract of $\chi 2282$ cells containing pDHFR7 over a 0.5 ml methotrexate-Sepharose affinity column. The extract was acidified to pH 5.8, passed over the column and the bound fraction was eluted with 2 mM folic acid in a 5 mM NaHCO₃ buffer at pH 8.5 as described elsewhere²⁸. Lane a, extracts of mouse cell line; lanes b and c, pDHFR 7; lane d, eluate from methotrexate column; lane e, pDHFR21; lane f, pBR322.



19 October 1978

Nature Vol. 275 19 October 1978

623

6 and unpublished the clone carrying stance, although the inserted in an orientation. This finding, and clone of protein that the mouse enzyme gh transcription into ter sequence located nase gene may occur. iminary data suggest- synthesised by cells ssor of Tp resistance n b (Fig. 7). Further quences in the distal le of serving as weak ains that extend into

t the bacterial clones and phenotypically A sequences: (1) the on shown by *in situ* mouse gene and by ode the amino acid zymatic activity and de sequences present actor, (3) the DHFR

bacterial cell extracts. un at constant current slab gel. The gel was et containing 0.15 M et with PBS. Peptides trips of a dry cellulose o anti-DHFR F(ab)₂ ed with antibody to I-labelled protein A. e filters were analysed ne d) was obtained by eg pDHFR7 over a 0.5 nn. The extract was and the bound fraction NaHCO₃ buffer at pH icts of mouse cell line; methotrexate column; BR322.

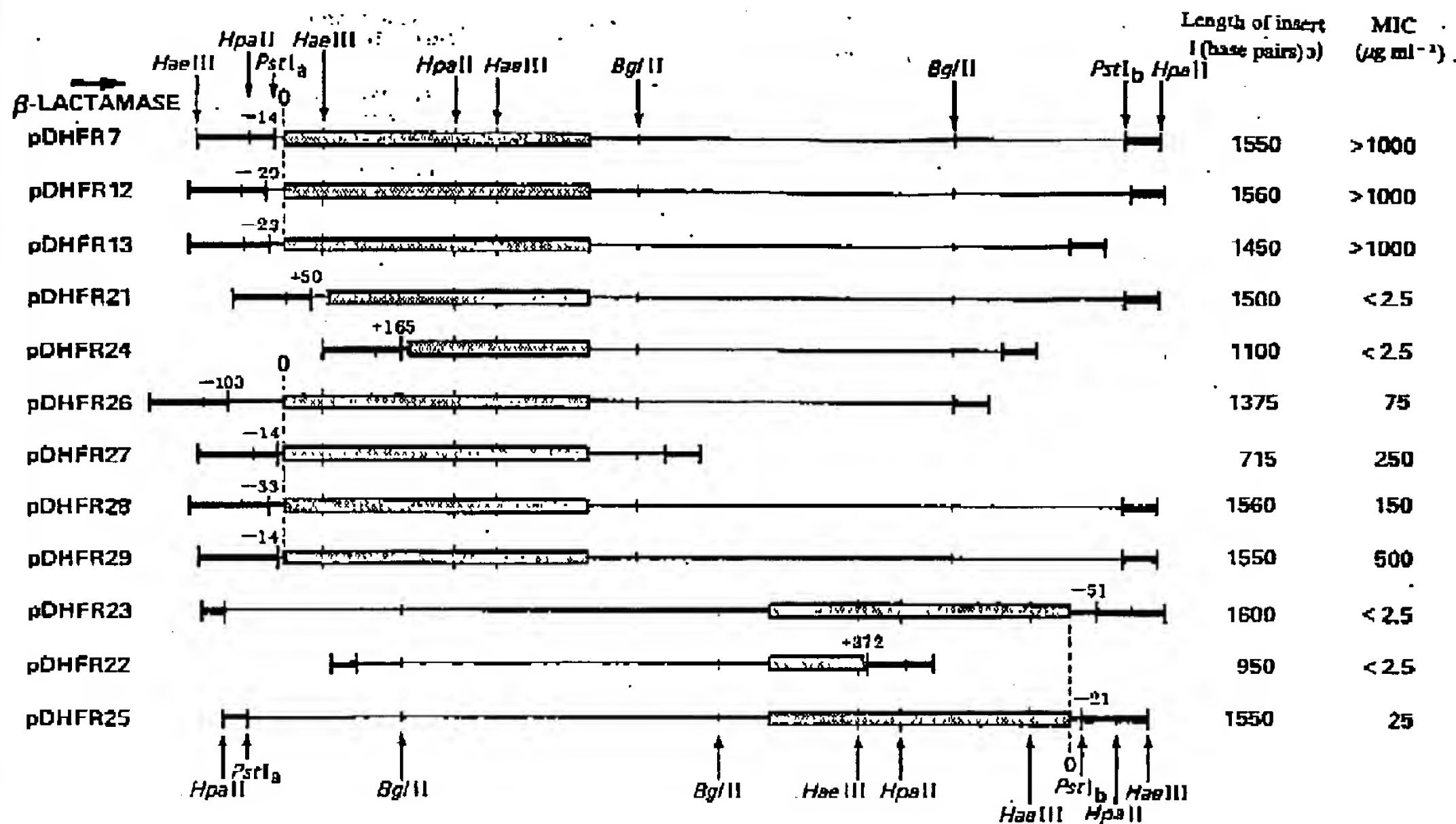


Fig. 7. Some structural and functional properties of chimeric plasmids and χ 2282 clones containing DHFR cDNA insert. The MIC of Tp for each of the χ clones was tested on M9 minimal agar plates containing biotin (2 μ g ml⁻¹), casamino acids (0.5%), DAP (50 μ g ml⁻¹) and Tp concentrations ranging from 0 to 1,000 μ g ml⁻¹. χ 2282 (pBR322) was used as a control and is sensitive to Tp at 2.5 μ g ml⁻¹, as determined by incubation at 32 °C for 3 d. Plasmid DNA isolated from each clone was digested with *Pst*I endonuclease¹⁵ at 37 °C for 3 h and extracted sequentially with phenol and ether. DNA was precipitated with ethanol, resuspended in 8 μ l TE buffer and electrophoresed in 1.2% agarose gels in TBE buffer to determine the length of the inserted fragment. ColEI plasmid DNA digested with *Hae*II endonuclease²⁷ was added to each sample as an internal molecular weight standard. Additionally, *Hind*III-generated fragments of SV40 DNA were used as an external standard²⁸. The standard error is \pm 100 base pairs with the method used. The orientation of the cDNA inserts in the vector plasmid was determined by gel analysis of plasmid DNA digested with *Bgl*II and *Hinc*II endonucleases; in addition, pDHFR25 was treated with the *Eco*RI and *Bgl*II enzymes to confirm the orientation of its insert. The direction of transcription of the β -lactamase gene of pBR322 is indicated by an arrow. The shaded area in each plasmid map indicates the structural sequence for DHFR. The numbers shown above each *Pst*I site indicate the distance (in base pairs) between the cleavage point and the first nucleotide of the DHFR structural sequence as described in Fig. 4. This distance was determined exactly by DNA sequence analysis for plasmids pDHFR 7, 12 and 13 and was estimated for the other plasmids by gel electrophoresis of fragments produced by either the *Hae*III or *Hpa*II endonucleases. For the chimeric plasmids that we subsequently sequenced, this estimate proved to be accurate within 5 base pairs. *Pst*I-*Bgl*II distances were determined for the chimeric plasmids by gel electrophoresis using *Hpa*II-cleaved pBR322 DNA as an internal standard. Nucleotides in the 5' direction on the mRNA from position 1 have negative numbers and are estimates obtained from gel analysis for all plasmids except pDHFR7 and 12.

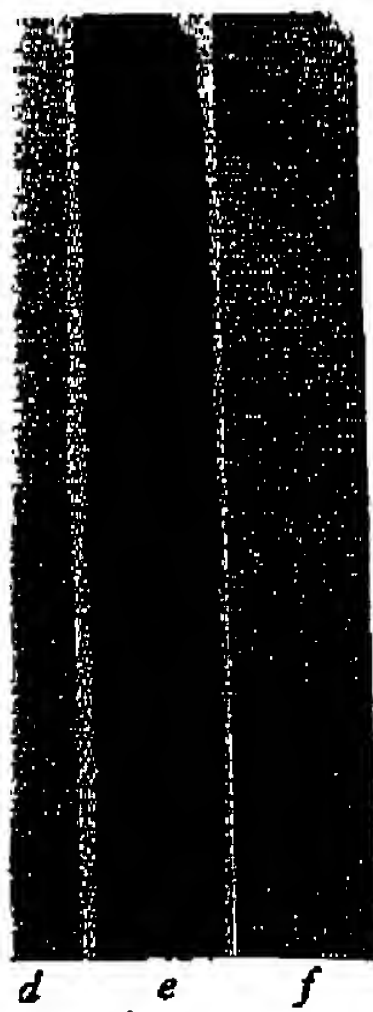
encoded by the constructed plasmids shows differential sensitivity to competitive inhibitors of DHFR characteristic of the mammalian gene product, and (4) the enzyme synthesised by bacteria containing DHFR cDNA is immunologically reactive with antibody made against mouse DHFR.

Note that the clones which express the highest levels of Tc resistance contain an immunologically reactive peptide having a size characteristic of the mammalian DHFR. A peptide of this size could potentially result from proteolytic cleavage of a fused β -lactamase-DHFR protein that was initiated at the β -lactamase ribosomal binding site. This proposal is consistent with the finding of large-sized protein species containing DHFR antigenic sites in extracts from cells that are either Tp^R or Tp^S (Fig. 6). More intriguing, however, is the possibility noted above that the *Pst*I-poly dG sequence constructed at the vector-cDNA junction can act together with the nearby ATG (AUG) translational start codon to bind mRNA to the bacterial ribosome and initiate DHFR peptide chains within the cDNA insert. If this interpretation is correct, initiation of peptide chains within other eukaryotic cDNA inserts may be obtainable in bacteria by use of the same structural relationships that have resulted in expression of the mouse DHFR coding sequence. Additional DNA

sequence analysis and investigation of the protein products encoded by chimeric plasmids should provide definitive information on this point and should help elucidate further the structural basis for the different levels of Tp resistance expressed by various clones.

Some of the bacterial clones we have isolated produce proteins that react immunologically with antibody to mouse DHFR but which are not biologically active. Our results suggest that an important obstacle to functional expression of mammalian DNA sequences in bacteria has been the development of an assay capable of detecting those clones that possess both a complete coding sequence and the correct nucleotide relationships to allow such expression. The strong phenotypic selection possible in the present experiments has provided an effective means of identifying and isolating expressing clones.

As the cloned coding sequence for mouse DHFR is selectable in higher organisms as well as in bacteria, it constitutes a powerful tool for the construction of eukaryotic cloning vectors, for the isolation of replication regions of eukaryotic chromosomal and extrachromosomal genomes (compare ref. 30), and for the isolation and characterisation of signals that control genetic transcription and translation in variety of species.



This work was supported by grants from the NIH, American Cancer Society and NSF to S.N.C. and R.T.S. H.A.E. is a Fellow of the American Cancer Society, California Division and acknowledges the hospitality and support of H. O. McDevitt. AMV reverse transcriptase was provided by Dr J. Beard, RNA by Drs F. Alt and R. Kellems, triazine derivative by Dr J. J. Burchall, and *E. coli* DNA polymerase by Dr H. Schaller. TdT was prepared by Dr R. L. Ratliff.

Received 4 September; accepted 15 September 1978.

1. Morrow, J. P. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **71**, 1743-1747 (1974).
2. Struhl, K., Cameron, J. R. & Davis, R. W. *Proc. natn. Acad. Sci. U.S.A.* **73**, 1471-1475 (1976).
3. Razkin, G. & Carbon, J. *Proc. natn. Acad. Sci. U.S.A.* **74**, 487-491 (1977).
4. Vopnek, D., Houtala, J. A., Jacobson, J. W., Giles, N. H. & Kushner, S. *Proc. natn. Acad. Sci. U.S.A.* **74**, 3508-3512 (1977).
5. Itakura, K. *et al.* *Science* **198**, 1056-1063 (1977).
6. Villa-Komaroff, L. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **75**, 3727-3731 (1978).
7. Burchall, J. J. & Flehling, G. H. *Molec. Pharmacol.* **1**, 126-136 (1965).
8. *Guidelines for Research Involving Recombinant DNA Molecules*. 23 June, (NIH, US Department of Health, Education and Welfare, 1976).
9. Curcio, R. III *et al.* in *Molecular Cloning of Recombinant DNA*, (eds Scott, W. A. & Werner, R.) 99-111 (Academic, New York, 1977).
10. Miller, J. H. in *Experiments in Molecular Genetics*, 218-220 (Cold Spring Harbor Laboratories, New York, 1972).
11. Grunstein, M. & Hogness, D. S. *Proc. natn. Acad. Sci. U.S.A.* **72**, 3961-3965 (1975).
12. Erlich, H. A., Cohen, S. N. & McDevitt, H. O. *Cell* **13**, 681-689 (1978).
13. Ali, F. W., Kellems, R. E., Berino, J. R. & Schimke, R. T. *J. biol. Chem.* **253**, 1357-1370 (1978).
14. Buell, G. N., Wickens, M. P., Fayat, F. & Schimke, R. T. *J. biol. Chem.* **253**, 2471-2482 (1978).
15. Smith, H., Blumberg, N. & Davis, J. *Nucleic Acids Res.* **3**, 343-353 (1976).
16. Kleinschmidt, A. K., Long, D., Jachertz, D. & Zahn, R. K. *Biochim. Biophys. Acta* **61**, 856-864 (1962).
17. Bolivar, F. *et al.* *Gene* **2**, 93-113 (1977).
18. Chang, A. C. Y. & Cohen, S. N. *J. Bact.* **134**, 1141-1156 (1978).
19. Enea, V., Vovis, G. F. & Zinder, M. D. *J. molec. Biol.* **96**, 495-509 (1975).
20. Cohen, S. N., Chang, A. C. Y. & Hsu, L. *Proc. natn. Acad. Sci. U.S.A.* **69**, 2110-2114 (1973).
21. Stone, D. & Phillips, A. W. *FEBS Lett.* **74**, 85-87 (1977).
22. Stein, J. A. & Steege, D. A. *J. molec. Biol.* **114**, 545-558 (1977).
23. Steege, D. A. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4163-4167 (1977).
24. Shine, J. & Dalgarno, L. *Proc. natn. Acad. Sci. U.S.A.* **71**, 1342-1346 (1974).
25. Burchall, J. J. *Ann. N. Y. Acad. Sci.* **186**, 143-152 (1971).
26. Baccantini, D., Phillips, A., Smith, S., Sinsky, D. & Burchall, J. J. *Biochemistry* **14**, 5267-5269 (1975).
27. Werkheiser, W. *J. biol. Chem.* **236**, 888-893 (1961).
28. Ali, F. W., Kellems, R. E. & Schimke, R. T. *J. biol. Chem.* **251**, 3063-3074 (1976).
29. Erlich, H. A., Levinson, J., Cohen, S. N. & McDevitt, H. O., in preparation.
30. Timmis, K., Cabello, F. & Cohen, S. N. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2242-2246 (1975).
31. Seeburg, P. H., Shine, J., Martial, J. A., Baxter, J. D. & Goodman, H. M. *Nature* **270**, 486-494 (1977).
32. Wickens, M. P., Buell, G. N. & Schimke, R. T. *J. biol. Chem.* **253**, 2483-2493 (1978).
33. Chang, L. M. S. & Bollum, F. J. *J. biol. Chem.* **246**, 909-916 (1971).
34. Wahl, G. M., Padgett, R. A. & Stark, G. R., in preparation.
35. Roychoudhury, R., Jay, E. & Wu, R. *Nucleic Acids Res.* **3**, 101-116 (1976).
36. Duggan, A., Boyer, H. W. & Goodman, H. M. *J. molec. Biol.* **96**, 171-184 (1975).
37. McDonnell, M. W., Shanon, M. N. & Studier, F. W. *J. molec. Biol.* **110**, 119-146 (1977).
38. Koppersock, Y. M. & Helinski, D. R. *Biochem. Biophys. Res. Commun.* **54**, 1451-1459 (1973).
39. Sharp, P. A., Sugden, B. & Sambrook, J. *Biochemistry* **12**, 3055-3063 (1974).
40. Maxam, A. M. & Gilbert, W. *Proc. natn. Acad. Sci. U.S.A.* **74**, 560-564 (1977).
41. Sanger, F. & Coulson, A. R. *FEBS Lett.* **87**, 107-110 (1978).
42. Kaufman, R. J., Datta, S. L. & Schimke, R. T., in preparation.
43. Lowry, O. H., Rosebrough, N. J., Farr, A. L. & Randall, R. J. *J. biol. Chem.* **193**, 265-275 (1951).
44. Oka, A. & Takahashi, M. *Nature* **264**, 193-196 (1976).
45. Fiers, W. *et al.* *Nature* **273**, 113-120 (1978).

letters to nature

N Galaxies—a new class of X-ray sources

BURBIDGE¹, classified extragalactic objects with bright nuclei into three classes: in order of increasing optical luminosity these are (1) Seyfert galaxies; (2) N galaxies; (3) quasars. Many Seyferts and several quasars have been shown to be X-ray sources. Here we show that N galaxies are also powerful X-ray sources. In fact, all six N galaxies in the 3C radio catalogue with redshifts less than 0.06 (ref. 2) are detected by the first full-sky survey of the Goddard Space Flight Center detectors (A2) on HEAO 1. X-ray emission has also been discovered from a strong Southern Hemisphere radio source, the N galaxy Pic A. Six of the seven objects are classified as broad-line radio galaxies (BLRGs)³⁻⁵ and one (3C371) is classified as a BL Lac object⁶. These objects are differentiated from the newly discovered class of 'emission line' galaxies⁷ by their high radio flux and extremely broad lines. Before HEAO 1 there was one confirmed N galaxy, 3C120 (ref. 8), and one suggested, 3C390.3 (refs 9, 10). We detect 3C120, strengthen the identification of 3C390.3, and present evidence for X-ray emission from four new sources, 3C111, 3C382, 3C371, and Pic A. We also suggest 3C445 as the identification of 2A2220-022 (ref. 11).

Figure 1 shows 90% confidence error boxes on the positions of these seven galaxies together with 4U (ref. 10) or 2A (ref. 11) error boxes where applicable. The 2-10 keV luminosity of these objects, as well as other relevant information, is given in Table 1. We have also looked for X-ray flux from the next closest N galaxies, PKS0521-36 and 3C227. No significant flux was seen, but the 90% upper limit on the X-ray luminosities (given in Table 2) are consistent with their luminosities being comparable to the seven detected N galaxies. Except for quasars, these N

galaxies are the most luminous class of compact X-ray sources yet detected.

Grandi and Osterbrock have classified radio galaxies spectroscopically as BLRG or narrow-line radio galaxies (NLRG) and have demonstrated a high degree of association between the classification 'broad-line' and the morphological type 'N'. In marked contrast to the situation for BLRGs, only 3C405 (Cyg A) of the five NLRGs listed by Grandi and Osterbrock, with redshifts less than 0.06, is a detected X-ray source. 3C317 is confused with Ab2052, a distance class 3 cluster, and so its luminosity is poorly determined. Upper limits on the flux from the other three NLRGs are given in Table 2. The deduced upper limits on the luminosities are considerably less than average X-ray luminosities for BLRG of $\sim 3 \times 10^{44}$ erg s⁻¹. Note that of the detected N galaxies, the one with the narrowest emission lines, 3C371, has the smallest X-ray luminosity.

The lack of X-ray emission from NLRGs contrasted to the virtual certainty of X-ray emission from BLRGs is similar to the distinction between type 1 and type 2 Seyferts, that is, none of the ~20 known X-ray emitting Seyfert galaxies are classical type 2 Seyferts. Grandi and Osterbrock have drawn attention to the optical spectral similarities and differences between BLRGs and type 1 Seyferts, and NLRGs and type 2 Seyferts, respectively.

The fact that many of these N galaxies possess compact radio components suggests that the synchrotron self-Compton process (SSC)¹² may be the principle cause of the X-ray emission. For 3C390.3, the best studied of these objects, a SSC model predicts a physical size to the X-ray emitting region of ~0.1 pc and a time scale for X-ray variability of months¹³. The detection of 3C390.3 by Uhuru and Copernicus at considerably higher flux levels and an Ariel 5 upper limit on 3C382 (ref. 14) ~20% lower than the present detection indicate possible variability in